Final report for project: R&D 7699-RB-01

The Preservation and the Decay of Military Skills

Contract N68171 - 95-C-9144

Avishai Henik, Ph.D.

Ben-Gurion University of the Negev

Beer-Sheva

Esther Brainin

Head of Research Section

Psychology Branch, Ground Corps Command

IDF

Varda Ze'evi

Head of Psychology Unit

Officers Training School

IDF

Dorit Schwarz

Ben-Gurion University of the Negev

Beer-Sheva

Address correspondence to:
Avishai Henik
Department of Behavioral Sciences
Ben-Gurion University of the Negev
Beer-Sheva, Israel 84105

e-mail: henik@bgumail.bgu.ac.il

20010502 084

Final report for project: R&D 7699-RB-01
The Preservation and the Decay of Military Skills

Avishai Henik, Esther Brainin, Varda Ze'evi, Dorit Schwarz

Introduction

Naive observations as well as academic studies of memory indicate that people forget what they have learned. Disuse can lead to skill deterioration, which, in the extreme, might mean that a skill is no longer functional when needed (Healy & Sinclair, 1996). Forgetting depends on a variety of factors, such as the quality of the skills required, the subject's characteristics, etc. Skill decay refers to the loss or decay of trained or acquired skill (or knowledge) after periods of nonuse. Skill decay is particularly salient and problematic in situations where individuals receive initial training of knowledge and skill that they may not be required to use or exercise for extended periods of time (Wisher, Sabol, Sukenik, & Kern, 1991).

Although various skill components may differ in their resistance to decay, in general, the longer the period of nonuse, the greater the decay. It is important to note that, although the length of the retention interval (i.e., the nonpractice period) has been cited as a powerful factor in retention, it is, albeit, a factor that may operate through mechanisms other than time per se (Arthur et al., 1998). Healy, Fendrich, Crutcher, Wittman, Gesi, Ericsson and Bourne (1992), in their comprehensive study of skill preservation, find that the retention rates for different skills vary considerably. Some skills are forgotten only minimally even after an extended period has passed since original

acquisition; for other skills there is a dramatic reduction of retention even after a short period.

The Israeli Defense Force (IDF) deploys its soldiers in units and expects that these soldiers' military skills, such as launching of anti-tank missiles or the aiming of a tank canon, be sufficiently high to allow their units to execute their tasks both in routine and in emergency situations. Various constraints, both operational and financial, determine the frequency at which reserve soldiers revitalize their skills, and in what manner. It is therefore important to examine the rate at which different military skills are lost, and what the determinants of the rate of the decay are.

During Operation Desert Storm a team of researchers of the ARI conducted a study of individual ready reserve (Wisher, Sabol, Sukenik, & Kern, 1991). The findings of this investigation point to a significant reduction of the military skills of these soldiers over a period of time. The researchers also point to a series of other studies revealing similar findings.

Factors affecting skill retention

Declarative vs. procedural knowledge

In Wisher et al.'s (1991) study, mentioned above, a number of factors that may affect memory were suggested. These factors included the soldier's initial level of knowledge (at the time of discharge from the military force) and the possibility of refreshing this knowledge during the period between discharge and call-up to a reserve unit.

Healy et al. (1992) suggested that variation in the decay rate might be due primarily to the nature of the skill, depending on whether the retention of procedures or the retention of facts is required. This distinction is based on

the distinction between procedural and declarative knowledge (Anderson, 1983). According to this distinction, declarative knowledge refers to the memory of facts and principles, whereas procedural knowledge refers to the remembering of processes. In terms of skill retention, procedural knowledge would be illustrated by the person's ability to perform the required task, even if he or she could not describe the procedures or the principles, which regulate performance verbally or in written forms. Healy et al. (1992) found that those skills which could be realized efficiently by memorizing procedures, decayed relatively little even after considerable time passed since the original learning.

Discrete vs. continuous tasks

Another distinction in the area of skilled performance is between closed-loop and open-loop tasks (named also as discrete versus continuous tasks, respectively). Closed-loop tasks, such as preflight checks and other fixed sequence tasks, usually involve discrete responses that have a definite beginning and end. On the other hand, open-looped tasks, such as tracking and problem solving, typically involve continuos movements or steps that are repeated and do not have a definite beginning and end (Arthur et al., 1998; Wisher et al., 1999). Studies show that open-loop tasks are better retained, even for extended time periods (months or years), than closed-loop tasks. Mengelkoch, Adams and Gainer (1971) used an aircraft simulator to train their subjects in various tasks needed for flight, from starting the engine and takeoff, to landing and shutdown of the engine. Two groups of subjects participated, with one group receiving twice the amount of the original learning time than the other. The retention interval was four months for both groups. The principal result was that discrete procedural response sequences had

significant loss over the retention interval, but proficiency in controlling flight parameters (tracking) showed no significant losses. Wisher et al. (1991) examining continuous and discrete tasks indicated similar results. The researches suggested that memory for perceptual motor skills, represented by accuracy scores at TOW gunnery, showed a resistance to decay, as was seen in recognition tests for school knowledge. Discrete, procedural skill, however, showed a large portion of decay. Marmie and Healy (1995) investigated the long-term retention of common military tasks. They studied tank gunnery using a TopGun tank simulator. The researchers referred to this task as a complex task because it involved discrete as well as continuous procedures, which required both motor and cognitive skills. In the TopGun subjects were required, for example, to locate the target through a continuous tracking procedure that had to be initiated by a discrete response (i.e., pressing and holding down a button). After locating the target, the subjects had to execute another continuous procedure to sight the target, followed by another discrete response procedure (i.e., pressing another button) to fire the target. The fact that the TopGun task contained discrete as well as continuous procedures lead the researchers to the prediction that overall retention would be hindered because the discrete procedures would be forgotten. The researchers trained six subjects in twelve sessions of 100 threat targets in 10 blocks. Each block included one target at each of ten possible ranges. Target location values were randomly assigned to the ten distinct occurrences of a given range value. The training period lasted for twelve sessions during four successive weeks, and the three retention test sessions occurred approximately 1, 6 and 22 months after the final training session. Accuracy of

shooting and time to identify and fire targets were measured. Results revealed that forgetting occurred only between the 1-month and 6-months retention tests, as was measured by the number of "kills" (i.e., the number of threats that were destroyed). Time to identify the target and time to fire did not show a significant change as a function of session number.

Individual differences
Level of original learning and over-learning

Besides the various task factors, it is possible to examine individual differences which affect skill decay and preservation. Researches examined whether the level of the original learning affected the ability of the trainee to use the knowledge and skills acquired in later circumstances.

"Original learning" refers to the amount of knowledge and skill the trainee has acquired by the end of training but before a job assignment. Practicing a task after it has been learned well enough to be performed correctly is referred to as "overlearning" (Wisher et al, 1999). Bahrick (1984) examined the retention of Spanish learned at school over fifty years. Multiple regression analysis showed that retention throughout a fifty year period is predictable on the basis of the level of original training. Large portions of the originally acquired information remain accessible for over fifty years in spite of the fact that the information is not used or rehearsed. This portion of the information, in a "permastored" state, is a function of the level of original learning, and the grades received in Spanish courses.

Overlearning provides additional training beyond that required for initial proficiency. Subsequently, a greater degree of learning is achieved. Hagman and Rose (1983) in their review present results of experiments that examined

retention of military tasks performed in an operational environment. Hagman (1980b, in Hagman & Rose, 1983) examined repetition effects in a task involving testing an alternator electrical output. They used four groups of participants that differed in the amount of times they performed the task during training (1, 2, 3, or 4 times). The results revealed that task repetition reduced task performance time and error on both the immediate and delayed retention tests. Additionally, time and error scores varied inversely with the number of repetitions performed. Schendel and Hagman (1980, in Hagman & Rose,1983) extended these findings by examining whether retention benefits could be gained by overtraining (i.e., performing additional repetitions after task proficiency was attained). They trained two groups of participants to disassemble and assemble the M60 machine gun. One group was trained to a proficiency criterion and the other group was trained to a mastery criterion. Retention rates, collected eight weeks after training, showed that the mastery group committed fewer errors than the proficiency group on the initial trial of retention, and needed fewer trials and committed fewer errors in relearning the task back to proficiency.

The findings presented above are consistent in showing that the level or degree of original learning is a significant factor in affecting forgetting. Overlearning may strengthen the bonds between stimuli and responses, and decrease the likelihood that responses will decay or be forgotten. Overlearning probably also enhances automaticity and subsequently, reduces the amount of concentrate effort demanded of the trainee (Arthur, Bennett, Stanush & McNelly, 1998).

Aptitude/ ability

Besides the level of original learning one can examine whether forgetting is affected by additional individual differences. Researches show that retention is higher for subjects with higher abilities (Farr, 1987).

Vineberg (1975, in Hagman & Rose, 1983) compared performance in 13 basic training tasks for soldiers in three mental categories defined by the Armed Forces Qualification Test (AFQT). He reported a direct relationship between mental category and overall task performance, both during baseline testing conducted at the end of training, and six weeks later. Performance decreased over the retention interval and the differences among groups were present at both testing sessions. These data suggested that differences in general ability level primarily affected the level of initial acquisition and that retention losses occured at the same rate regardless of ability. Arthur and his colleagues (Arthur, Bennett, Stanush & McNelly, 1998) applied meta-analytic techniques to find the factors that had the major influence on skill decay and retention. They suggested that the role of individual differences in skill decay or retention is often confounded by the degree of original learning. Specifically, although results consistently find that higher ability individuals retain more knowledge or skill over periods of nonuse than lower ability individuals, it is argued that higher ability trainees really acquire more knowledge, skill, or both, in the same amount of time than lower ability trainees. There is some evidence, however, that lower ability learners forget a larger portion of abstract, theoretical material than do higher ability individuals (Arthur et al., 1998).

The current study

The current study was designed as a prototype for investigating the preservation of military skills. We concentrated on individual skills, which aside from their own importance, also have a significant effect on team performance (Spector, 1992).

The selection of skills that were examined was based on the following considerations: (a) that they were considered to be sufficiently important by the IDF to warrant such analysis; (b) that it would be relatively simple to asses the level of performance regarding these skills; (c) that they be sufficiently heterogeneous to allow for a relatively wide variation in task characteristics and still be similar enough to examine them jointly. We focused on aiming and shooting skills on a variety of weapons such as the tank main-gun and *Dragon* and *Tow* missiles.

Skills and their testing

Targeting and shooting skills involve a number of rules or guiding principles, which control performance while carrying out the skill. Accordingly, we examined the skill specifically related to shooting, which reflects mainly procedural components and knowledge components of this skill. The knowledge components probably reflect mainly declarative knowledge. We studied both aspects, since their preservation rates may have differed.

The selected skills are similar in that they all reflect components regarding the aiming and firing of weapons. However, there are also qualitative differences between these skills (see below).

Tank gunner

The tank gunner is not charged with making the decisions about firing the shell, but only for carrying them out. That is, he does not determine the

type of ammunition to be used, the target, the firing method (in most cases), or the corrections. All these decisions are made by the computer or by the commander. During the shooting the gunner is required to enter the type of shell into the computer and measure the firing distances with a mini-laser (this measurement is performed by first placing the cross on the target and then starting the mini-laser). Since the cross jumps after operating the laser, the gunner had to place it on the target again, check the data in the computer, and only then does the actual firing take place. If the target is missed, the gunner is supposed to correct his aim according to the tank commander's directives (by pushing the button, or by correcting the position of the cross), return the cross to the aiming point, and pull the trigger. The gunner must remember a number of principles concerning how to aim at the target, and the techniques for firing during an emergency (system failure) and for firing under special conditions such as a moving target, shooting while in motion, etc. The tank gunner can rely on his commander for assistance in fostering his recollection of the relevant tasks. In practice, the tank commander will almost always come to the tank gunner's aid, correcting the latter's failure to recall any given procedure. In this case, one of the important attributes of firing is that after pulling the trigger, the gunner has no control over the trajectory of the shell; at this point, the firing episode has, in essence, been completed. It should also be noted that each shooting episode consists of a series of four shells. In other words, satisfactory performance is not necessarily judged on the basis of the first shell, but as a function of a hit among a series of shots.

Tow and Dragon

The aiming of a missile differs from that of the canon of a tank in that the former is equipped with the capability to guide the missile to its target. The missile can be steered even after it has been launched. The firing technique involves the placement of the homing cross on the target, launching of the missile, and keeping the cross on the target, whether stationary or moving, until it is impacted. Aiming of both the Dragon and the Tow requires considerable motor skills. Here, we shall refer to each missile launch as a separate event, as the missile is steered en route to the target by the soldier firing it.

In testing the various aspects of the skill involved, it is interesting to examine the differences between tank gunners and missile launchers. The latter engage in routing the missile toward the target, so that it is possible to directly observe deviations, or the amount of time on-target, from the instant the missile is launched until the target is hit. Obviously, since there are considerable differences in the skill required in these two cases (tank gunners and missile launchers), any differences in the pattern of skill preservation among the two groups should be regarded with caution.

The intention of the current study was to examine the mentioned skills with respect to the following points:

1. The rate of the skill decay.

2. The impact of skill refreshing on the rate of skill decay.

3. The possibility of affecting procedural knowledge by declarative knowledge, and vice versa.

4. The impact of individual differences and degree of original learning on the rate of skill decay.

Method

Subjects

      We tested a total of 330 subjects, who were released from the army during the periods of July 1996 to August 1999. The final sample included 204 subjects (see Table 1). The sample was composed of 82 Tank Gunners, 79 Tow operators, and 43 Dragon operators. The rest of the subjects (126 participants) were either not available after discharge or did not participate in the final testing session.

      When first tested, all subjects were in the last two months of their military service, which commonly lasts three years.

Table 1. Distribution of subjects according to experimental groups.

	Tank Gunners	Tow	Dragon	Total
At discharge	111	138	81	330
Group A (After 6 and 18 months)	30	18	12	60
Group B (After 12 and 18 months)	14	23	12	49
Group C (After 18 months)	38	38	19	95

Simulators

      Tanks- The IDF employs Rafael simulators for the tanks' main gun. In the simulator the gunner is presented with a topographical display on a TV screen. For each event the gunner is required to identify the territorial display,

to place the cross on the target, and operate the mini-laser and shoot if the firing is done from a stationary position, or just shoot if the firing is done in motion. Thereafter, the gunner is required to make the necessary corrections, if any, according to the directives he receives, until the target is hit.

Missiles (Tow and Dragon)- The IDF employs Simtec simulators. Like the Tank gunners' simulators, these simulators present the missile launcher with a topographical display on a TV screen. For each event the launcher identifies the territorial display, places the cross on the target and launches the missile. In case of a moving target, he launches the missile only after target acquisition. Thereafter, he is required to maintain the cross on the target continuously, until the target is hit.

Tests

There were three types of tests: practical, knowledge and memory tests.

Practical tests

Tests were conducted individually. Professional military instructors of the particular simulator (Tank main gun, Tow or Dragon) tested the soldiers. The instructors were asked to refrain from helping the soldiers at any point of the test. They only gave the very basic commands, regularly given during operation by commanders in the real time.

Tank main gun- There were twelve scenarios, of which six were at night. Each scenario required a different shooting technique: firing from a stationary position, a consecutive sequence, firing in motion, firing on a moving target, and *Tefa* (correcting according to lens-self-correction). Every soldier performed the day scenarios first and the night scenarios second.

Within each type (day or night), scenarios were randomly presented. One random order, that was initially randomize, was employed for all soldiers.

Every shooting series was composed of four shells. Firing was completed either by a hit or at the end of the series. The simulator enabled us to enforce a certain error for the first shell (the first shell could land at different locations, such as: (a) long and wide to left; (b) short and wide to right; (c) questionably right). Each of these error types was allocated to a different scenario. Introducing errors in this way let us look at the soldiers' ability to correct mistakes during shooting. Running time for each individual was 25 minutes.

Missile Launchers- Choice of scenarios involved input from commanding officers of the Tow and Dragon units. To choose scenarios we employed the following procedure. First, we listed the parameters according to which the various scenarios could be sorted. These were as follows: range of target, moving vs. steady target, visual conditions (e.g., daylight vs. night). In addition, if the target was moving the scenarios could be with fast or slow movement.

Tow-There were twelve scenarios, of which half were night scenarios and half were day scenarios. Scenarios were organized from easy to hard, except for the hardest scenario, which appeared third in the list. The latter was done in order to enable us to test for difficulties with high level scenarios not confounded by practice. During the scenario the soldier sat on a chair in front of the missile launcher. Running time for each individual was 25 minutes.

Dragon- There were ten scenarios, all day scenarios (the available night scenarios did not fit night circumstances). Similar to the Tow, scenarios

were organized from easy to hard except for the hardest scenario which appeared third. During the scenario, the soldiers knelt down on his knees and hold the launcher on one of his shoulders. Unlike the Tow, this operation required investment of force, and control of movement and breathing (one should not move one's body while aiming and shooting). In addition, once the missile was launched, there was an immediate reduction in the weight of the launcher (this was achieved by releasing a six kg weight from the rear side of the launcher). The abrupt reduction in weight has to be taken into account by the operator so that he is able to keep the launcher still.

There was a need to calibrate the simulator for each individual, before starting the test. In addition, loading scenarios from the disc to the system took a few minutes. In light of these time consuming operations (calibration and loading), and because the various scenarios were on two different laser discs, we used two separate simulators in parallel. To keep the predefined order of scenarios, soldiers moved from one simulator to the other, as needed. Running time for each individual was 25 minutes.

Knowledge tests

Officers in charge of the various military skills under study, composed the tests. The tests included questions regarding firing techniques (e.g., visualizing the target through the lens during shooting), firing rules (e.g., initializing the system), matching ammunition to type of target, and limitations of the systems (e.g., what are the circumstances for reinitialization of the system). Most questions were multiple-choice type. There was no time limit and the mean time for completing the test was 25 minutes.

Tests for the Tank gunners were designed to encompass all firing techniques. There were between 4 to 6 questions for each technique. Because there were four different types of tanks, the knowledge test covered themes that were shared by all tank gunners.

Memory tests

Tests were designed to fulfill the following criteria: a) encompass the relevant type of memory; b) be short and not too complicated; c) items in the tests were not to be related to the military; and d) there were to be paper and pencil tests. There were five memory tests: a) Verbal free recall - The test was composed of 15 words printed on a card. The words were exposed for one minute followed by 30 seconds of a distractor task. The distractor task consisted of counting backward by 3's from the number the experimenter presented at the end of the one-minute exposure (learning) period. During the test phase, soldiers were asked to write down all items recalled, with no time limit. The test score was the number of correct answers. b) Visual free recall - The test was composed of 15 line drawings of standard objects on a card, e.g., a car, a bottle, or a ball. During the test phase subjects were asked to write down the names of the objects. Other features of this test were the same as in the verbal free recall test. c) Paired associates - The test was composed of 15 pairs of unrelated words printed on a card. The words were exposed for one and a half minutes. During the test phase soldiers were exposed to a list of words, each of which could be the first or the second member of the pair. Soldiers were asked to recall (and write down) the other member of the pair. Other features of this test were the same as in the verbal free recall test. d) Semantic memory - Subjects were asked to retrieve (and write down) as many

items as possible within 45 seconds, from a given category. The experimenters presented a category name (fruits, cars, or occupation) and started the stopwatch. At the end of the time period the experimenter asked the subject to stop. For each category, we counted the number of items produced. The score of a given individual was the average of these numbers.

e) Memory for order - Subjects were exposed to a list of twelve words presented from top to bottom on a page. They were asked to remember the order of the items (from top to bottom). At the end of the presentation phase, subjects received 30 seconds of a distractor task (as in the three of the memory tests). In the test phase subjects were given twelve words on small cards and asked to arrange them according to the earlier order of presentation. The score was based on a correlation between the order of the original list and the order produced by the subject.

Procedure

Tank main gun tests were conducted at four field bases. The IDF employs four different simulators, with similar functioning, for tank main guns. The training base simulator is used throughout the year for different tanks. We found it impossible to match the research schedule with the training base schedule, thus, it was easier to conduct the tests in the simulators of the field units. All tank simulators functioned the same, so that the amount of heterogeneity due to the above testing at different bases seemed to be minor. Tow and Dragon tests were conducted at a central training base.

Soldiers were invited to a 'revitalization day'. One member of the research staff met soldiers in a waiting room as they arrived in the base. Soldiers were then introduced to the study and its goals. These introductory

comments were presented to groups of soldiers or to individual soldiers according to their availability. Before taking any test, soldiers filled out a 'general details' questionnaire. This questionnaire included personal details, questions regarding usage of the missile launcher or tank main gun during their military service, and participation in professional courses during their military service. Soldiers were randomly allocated to two groups. One group received the practical tests first and the knowledge test second. The other group received the knowledge test first and the practical tests second. Memory tests were conducted while subjects were waiting for practical or knowledge tests. This practice was adopted because memory tests were not supposed to be influenced by the practical/knowledge tests that preceded them. At the end of the day each soldier was asked to fill out a final questionnaire. In the final questionnaire soldiers assessed their performance in the various tests.

Measures

Dependent variables: measurements were taken of both the practical and the theoretical components of each of the skills.

Practical: in the course of each scenario, both the *speed* and the *accuracy* of performance were assessed. Speed was measured based on the launching time (in seconds), i.e., the amount of time elapsed between the firing command and the launch.

The accuracy of performance was assessed using a number of measures: 1) the proportion of time on the target, 2) vertical and horizontal aiming deviations from target (for each launch, vertical and horizontal

standard deviations between the center of the cross and the ideal or shortest launcher-target line were computed). 3) hit versus miss.

Knowledge: the knowledge measure was the proportion of the correct answers in the written test given to the missile launchers and the tank gunners.

Independent variables: the independent variables of this study were those which were hypothesized to affect skill preservation. Before presenting the variables we would like to describe the chronological aspect of this study, which is delineated in the table below. The plus sign (+) signifies the testing of skill, as described above.

Table 2. Experimental group testing periods

Testing period	Discharge	6 Months	12 Months	18 months
Group A	+	+		+
Group B	+		+	+
Group C	+			+

All three groups were tested immediately before discharge from regular military service, and again 18 months after discharge. Groups A and B were tested six and twelve months after their discharge, respectively.

The final test which was scheduled for all three groups at the conclusion of the study after 18 month facilitated analysis of several additional aspects. The comparison of groups A and B with group C enabled us to examine whether the utilization of a simulator (the testing of various components of the skill constituted training in itself) restored relevant knowledge, and to what extent. Moreover, the comparison of group A and B

answered the question whether the timing of refreshing of the skill - after six or after twelve months - affected its effectiveness.

As noted, each test included a practical and a knowledge segment. One half of the subjects took the practical tests first, while the other half started with the knowledge part. By varying the order of these two components of the test, it was possible to check whether the recall of practical skills can be improved by a refreshing of theoretical knowledge, and vice versa.

Accordingly, the independent variables in this study were as follows:

Time of recall - the amount of time elapsed between the first and the second test. Three levels were used: 6, 12, and 18 months.

Revitalization - whether or not subjects had an opportunity to refresh their memory. This corresponds to the comparison between the performance of groups A and B with group C after 18 months.

Timing of revitalization - the relevant phases at which the skill was refreshed. This corresponds to the comparison between the performance of groups A and B after 18 months.

Order of testing - whether the practical tests reinforced knowledge, or vice versa.

Measures of motivation

Actual performance depends not only on the level of skill at the time of measurement, but also on the motivation at that time. Therefore, in order to control for potential decrement in motivation as a function of the retention interval, we measured (through a questionnaire) the motivation of the participants engaging in performance.

Because of some problems in administering these questionnaires we collected only a limited number of such questionnaires.

Aptitude measurement

At the time soldiers enlist in the IDF they undergo various qualification tests. For each participant in the study we obtained the DAPAR (IDF equivalent of the AFQT) and KABA (IDF measure composed of DAPAR, grade in personality interview, education and Hebrew levels), and employed them to explore whether estimates of skill retention vary as a function of aptitude level.

Level of original learning

Missile launchers and tank gunners engage in qualification courses during their military service. For each participant in the study we obtained the final score achieved at the end of the qualification. These scores were employed to explore whether estimates of skill retention vary as a function of the level of original learning.

Results

This section is organized by the various military skills and issues discussed in the introduction. The various issues are presented for each military skill separately. At the end of this section we present results across weapon systems. We start with Tank gunners, followed by Tow and Dragon operators.

In the following figures we present changes in performance as a function of time elapsed since discharge. The values for each point in each figure are based on different numbers of subjects. The relevant numbers of

subjects (N) are presented on the last three lines of Table 1 (6 months, 12

months, and 18 months after discharge, respectively).

Tank gunners

1. Skill decay

Knowledge test - Figure 1 depicts changes in knowledge test scores with time

since discharge.

**Figure 1. Changes in knowledge tests as a function of time since discharge**
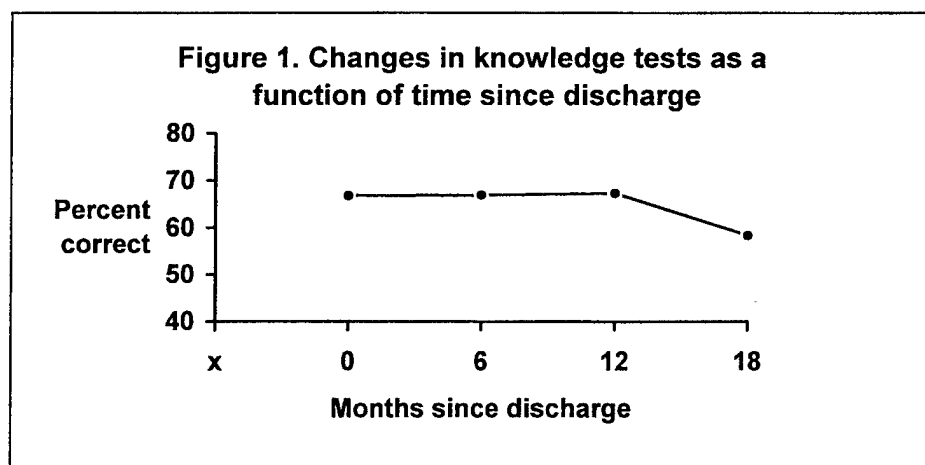
A one-way ANOVA showed a significant reduction in performance with

change in time since discharge [$F(3,15) = 4.59$, $p < 0.05$].

Practical performance - The following two figures present data regarding the

number of hits, vertical deviations, and horizontal deviations from targets.

The range for the number of hits is between 0 and 14. As can be seen

in Figure 2, the number of hits declines with time since discharge [$F(3,159) =$

3.69, $p < 0.05$].

Figure 2. Changes in hits as a function of time since discharge

N of hits

Months since discharge
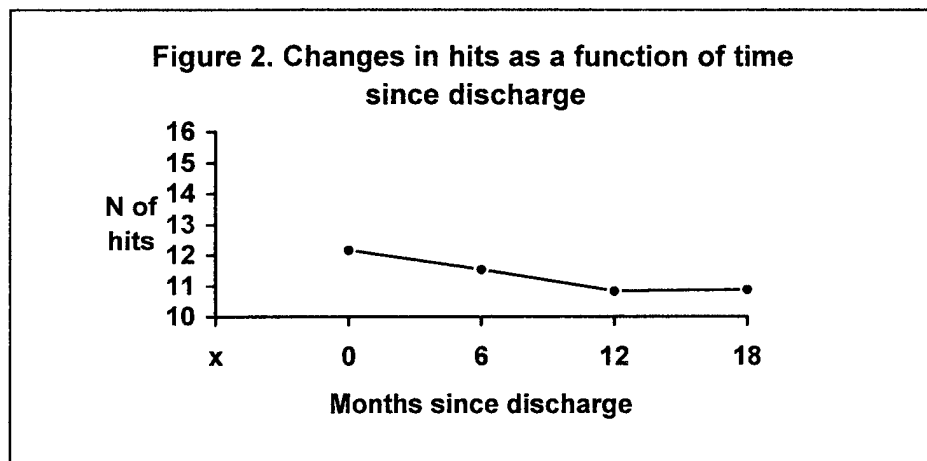
The next figure presents results pertinent to accuracy in shooting. It presents horizontal and vertical deviations from target locations. For each subject we computed the mean deviation in mills (one mill is 1/1000 of a degree). These means were entered into two separate one-way ANOVAs for the horizontal and vertical deviations: for the horizontal deviations [$F(3,183)$ =4.69, $p< 0.01$], and for the vertical deviations [$F(3,179) = 2.71$, $p< 0.05$]. As can be seen in the figure, deviations increased as a result of time since discharge.

Figure 3. Changes in deviations as a function of time since discharge

Deviations

—•— Horizontal
—■— Vertical

Months since discharge
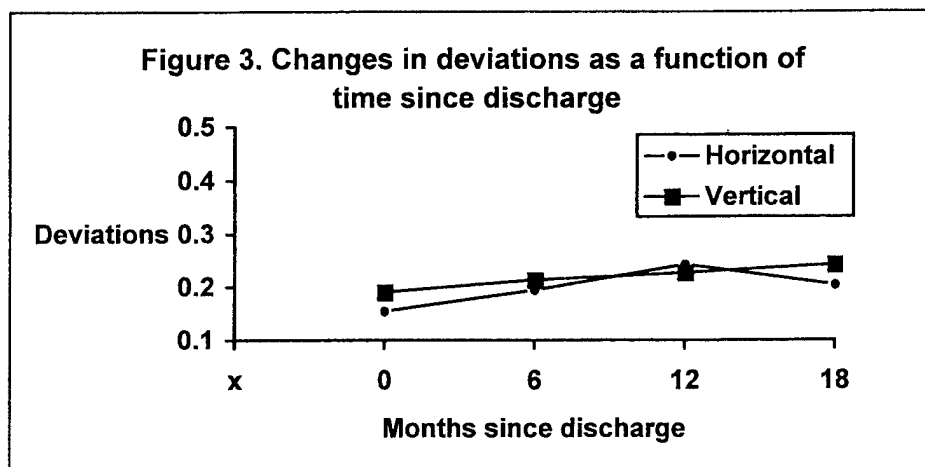
2. Effects of refreshing

In order to examine effects of revitalization on performance, we compared performance at 18 months since discharge for three groups of subjects, i.e., those who experienced revitalization (see Table 1, lines 2 & 3; subjects who were tested at 6 or 12 months after discharge) with subjects who had no military practice between discharge and the test at 18 months (see Table 1, line 4).

Knowledge test – The comparison of the three groups of subjects showed a significant difference [F (2,78) =4.38, p<0.05].

**Figure 4. Changes in knowledge tests as a function of revitalization**
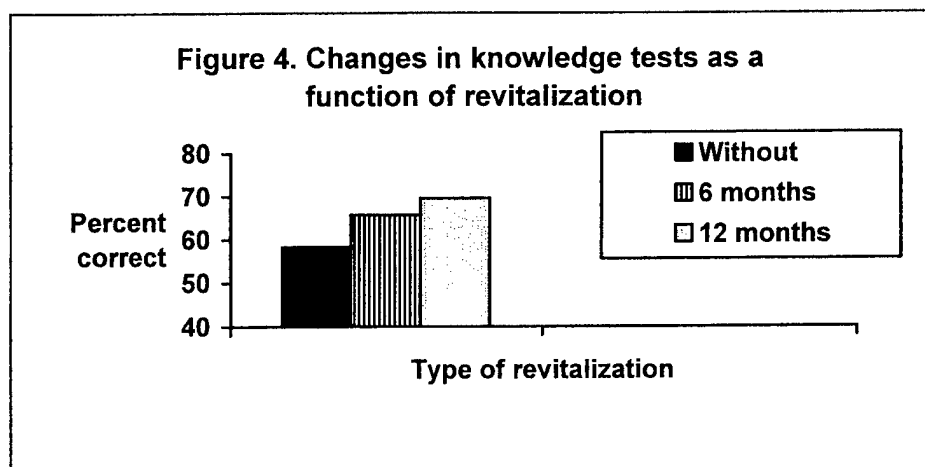
Additionally, we contrasted differences between subjects who experienced revitalization (i.e., either after six or twelve months since discharge) with subjects who had no revitalization. This comparison turned out to be significant also [F(1,78)=8.72, p<0.01]. However, there was no difference in knowledge after 18 months as a function of the timing of the revitalization, i.e., after 6 or 12 months [F(1,78)<1, ns].

Practical performance - Figure 5 depicts changes in the number of hits as a function of revitalization. It seems that there was no significant difference with revitalization [F(2,74)=1.83, ns].

**Figure 5. Changes in hits
as a function of revitalization**

N of hits

- ■ Without
- ▥ 6 months
- ☐ 12 months

Type of revitalization
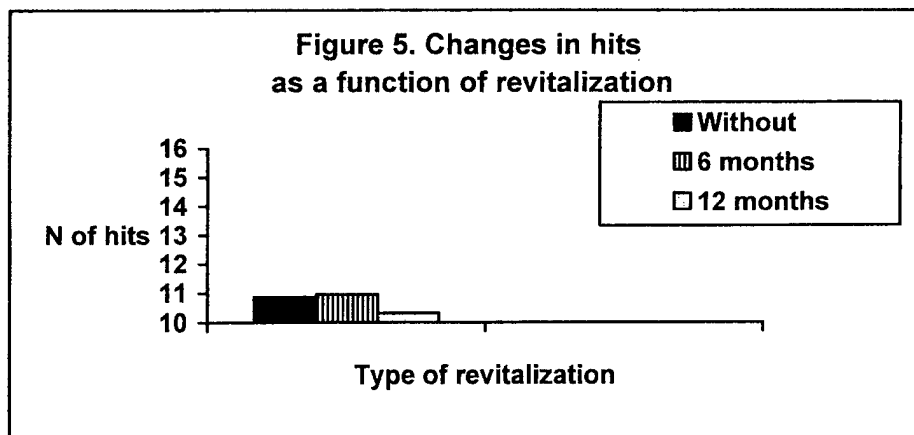
The next figure depicts changes in deviations from the target as a function of revitalization and type of deviation. It seems that for both types of deviations there was no significant drop with revitalization. Mean deviations were entered into two separate one-way ANOVAs for the horizontal and vertical deviations: for the horizontal deviations [$F(2,78) = 0.69$, ns], and for the vertical deviations [$F(2,74) = 0.97$, ns].

Figure 6. Changes in deviations as a function of revitalization and type of deviation

Deviations

- ■ Without
- ▥ 6 months
- ☐ 12 months

Horizontal   Vertical

Type of deviation
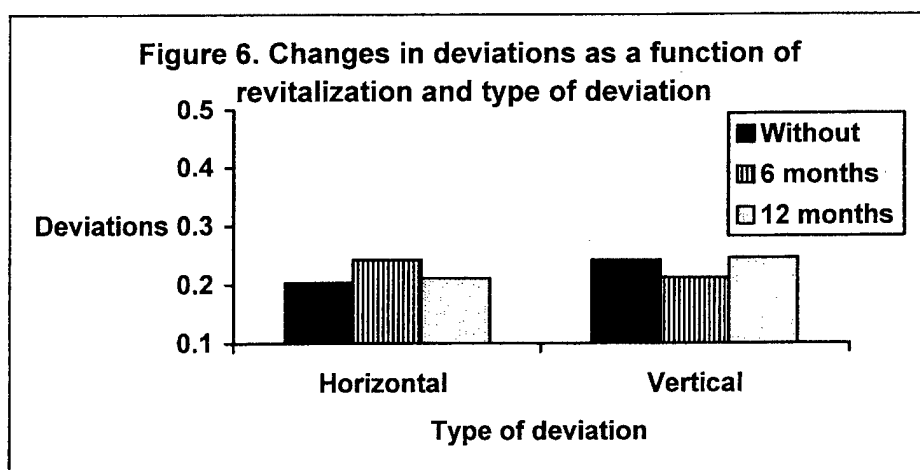
3. Relationship between knowledge and practice

Half of the soldiers participated in the knowledge test first and the practical test second and half participated in the tests in the reverse order.

The order of tests had no effect on performance (all F values were smaller than 1).

Tow operators

1. Skill decay

Knowledge test - Figure 7 depicts changes in knowledge test scores with time since discharge. There was a significant decline in percent correct with time since discharge [$F(3,240) = 14.09$, $p < 0.001$].

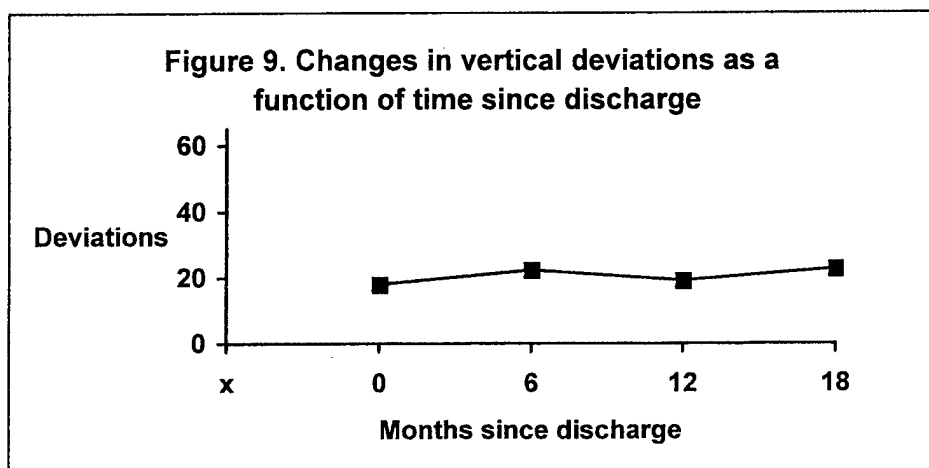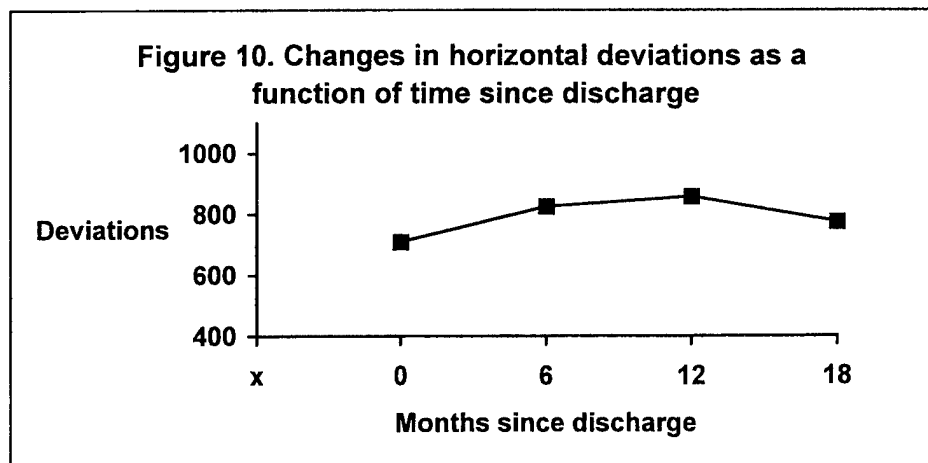Figure 7. Changes in knowledge tests as a function of time since discharge
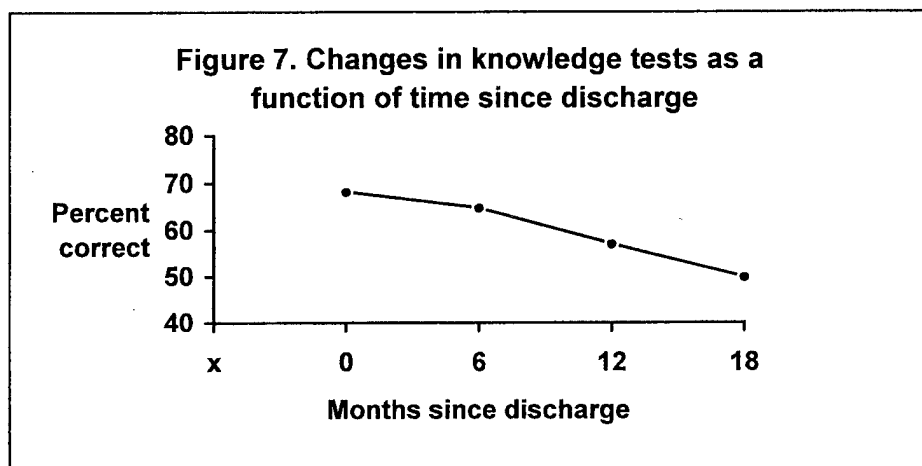
Practical performance - The following four figures present data regarding proportion of hits, vertical deviations and horizontal deviations from targets, and time on target.

As can be seen in Figure 8, the proportion of hits declines with time since discharge. A one-way ANOVA showed a significant effect [$F(3,181) = 2.81$, $p < 0.05$].

Figure 8. Changes in proportion of hits as a function of time since discharge
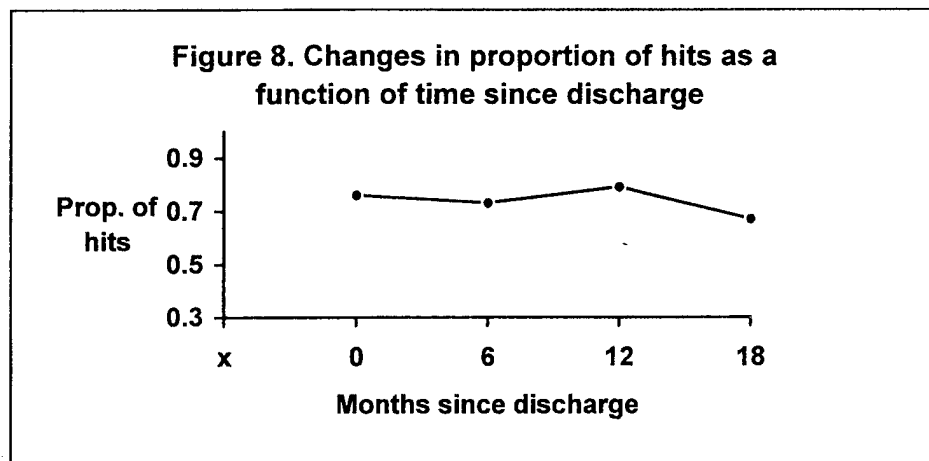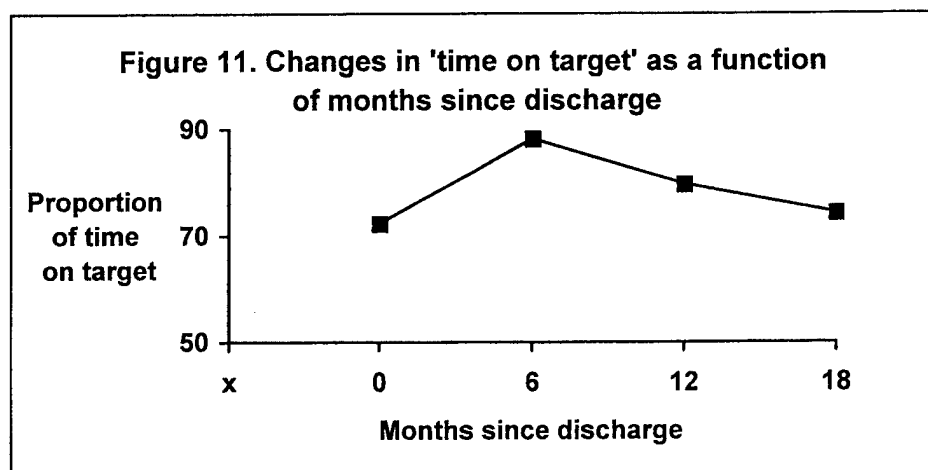

The next figures present results pertinent to accuracy in shooting. Figures 9 and 10 present vertical and horizontal deviations from target locations, respectively. For each subject, horizontal and vertical deviations were computed for each scenario. These deviations were used to compute a mean deviation for each subject across scenarios. These deviations were entered into two one-way ANOVAs. The ANOVA of the vertical deviations showed a non significant result [$F(3,239) =1.80$, ns], and the ANOVA of the horizontal deviations showed a non significant result as well [$F(3,238) = 1.20$, ns].

Figure 9. Changes in vertical deviations as a function of time since discharge

Figure 10. Changes in horizontal deviations as a function of time since discharge

Devotions (y-axis): 400, 600, 800, 1000

Months since discharge (x-axis): 0, 6, 12, 18

The simulator produces a score for 'time on target'. This is the proportion of time that the missile was within the target range, out of the total time since launching the missile. The results are depicted in Figure 11. An ANOVA produced a significant difference between the various times of testing (i.e., months since discharge), $[F(3,221) = 6.12, p < 0.01]$. The figure shows that at discharge soldiers did not show very good performance but later on their performance improved. From 6 months after discharge there is a decline in the ability to keep the missile on target. This decline was significant $[F(1,221) = 9.594, p < 0.01]$.

Figure 11. Changes in 'time on target' as a function of months since discharge

Proportion of time on target (y-axis): 50, 70, 90

Months since discharge (x-axis): 0, 6, 12, 18

2. Effects of refreshing

<u>Knowledge test</u>- Figure 12 depicts changes in knowledge test scores as a function of revitalization. The comparison of the three groups of subjects showed a significant difference [$F(2,77) = 3.85$, $p<0.05$]. We contrasted subjects who experienced revitalization with subjects who had no revitalization and found a significant difference [$F(1,77)=5.80$, $p<0.05$]. In contrast, there was no difference in the knowledge test after 18 months as a function of the timing of the revitalization [$F(1,77)=1.89$, ns].

Figure 12. Changes in knowledge tests as a function of revitalization
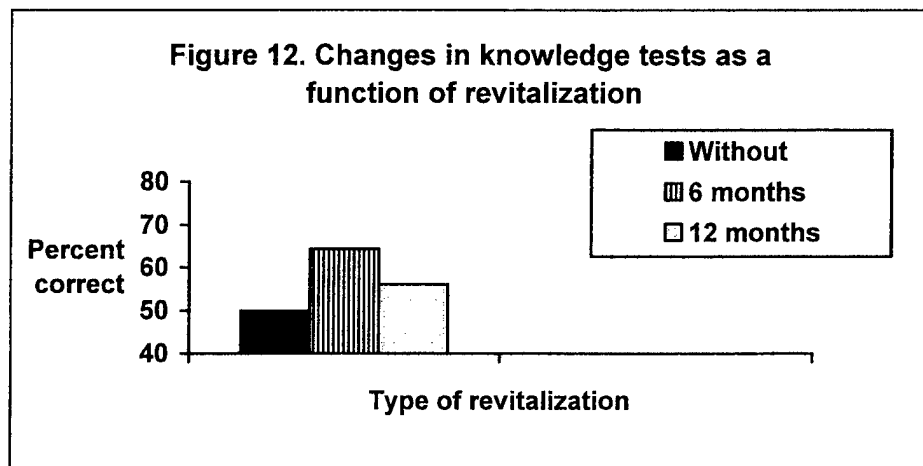
<u>Practical performance</u> - As before, we examined the effect of revitalization on performance. The comparisons of the proportion of hits were not significant [$F< 1$] and the same was true for the horizontal and vertical deviations (both $F<1$). There was no effect of revitalization on time on target, as well [$F(2,69)=1.93$, ns].

3. Relationship between knowledge and practice

The order of tests had no effect on performance (all F values were smaller than 1).

Dragon operators

1.Skill decay

Knowledge test – Figure 13 depicts changes in knowledge test scores with time since discharge. There seems to be a significant decline in percent correct with time since discharge [$F(3,138) = 6.18$, $p < 0.01$].

Figure 13. Changes in knowledge tests as a function of time since discharge
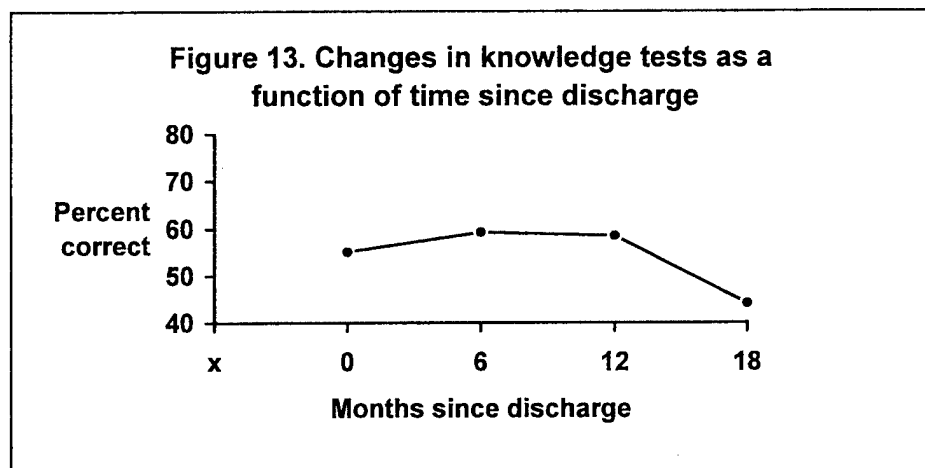
Practical performance - The following figures present data regarding proportion of hits, and time on target.

The range for the proportion of hits is between 0 and 1. As can be seen in Figure 14, there was no significant decline in proportion of hits as a function of time since discharge [$F(3,132 = 1.14$, ns].

Figure 14. Changes in proportion of hits as a function of time since discharge
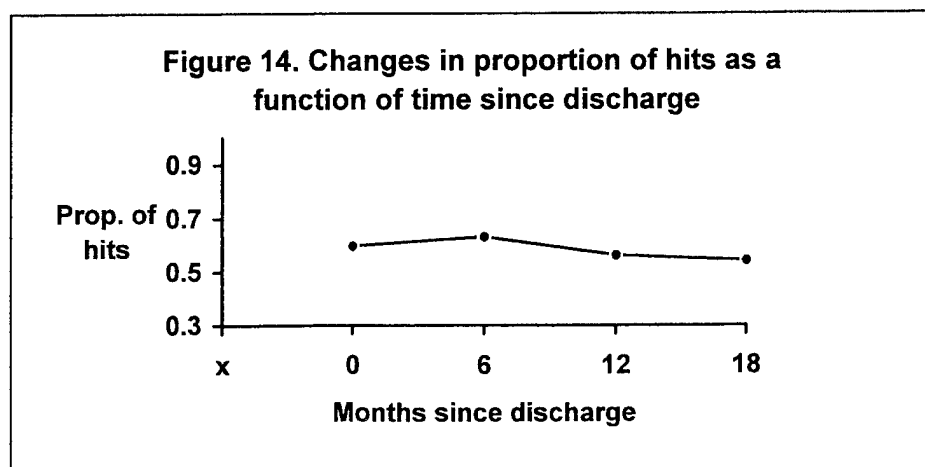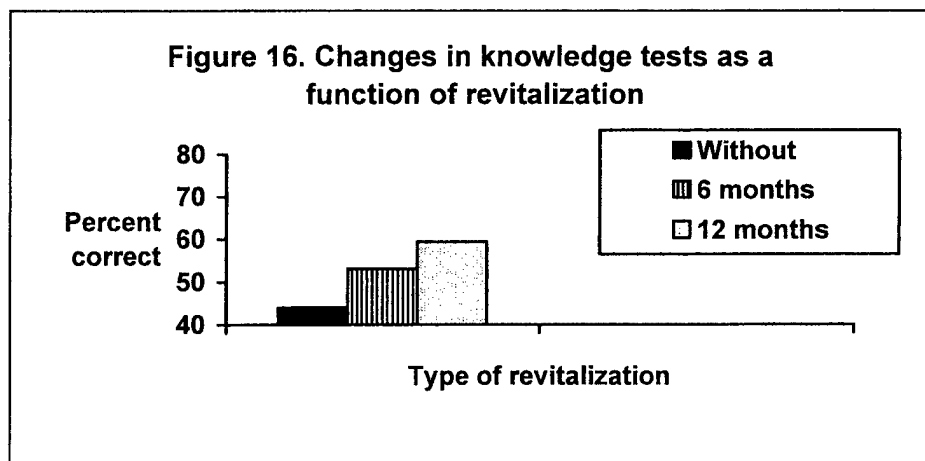
Like the Tow simulator, the Dragon simulator produces a score for 'time on target'. The results are depicted in Figure 15. An ANOVA produced a significant difference between the various times of testing [$F(3,133) = 3.82$, p< 0.05].

Figure 15. Changes in 'time on target' as a function of months since discharge
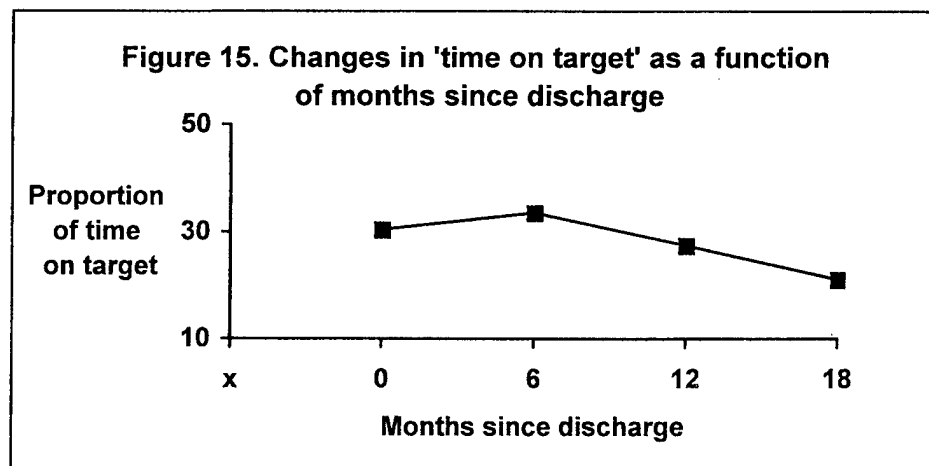
2. Effects of refreshing

Knowledge test- Figure 16 depicts changes in knowledge test scores as a function of revitalization. The comparison of the three groups of subjects showed a significant difference [$F (2,40) = 5.05$, p<0.05]. As before, there is a significant difference between subjects who experienced revitalization and subjects who had no revitalization [$F(1,40)=8.79$, p<0.01], however, there was no difference as a function of the timing of the revitalization [$F(1,40)=1.31$, ns].

Figure 16. Changes in knowledge tests as a function of revitalization

Percent correct

80
70
60
50
40

■ Without
▥ 6 months
☐ 12 months

Type of revitalization

Practical Performance – Figure 17 depicts the effect of revitalization on proportion of hits. The comparison of subjects who experienced revitalization with subjects who had no military practice between discharge and the test showed a marginally significant difference [F (2,40) = 3.04, p<0.06].

Figure 17. Changes in proportion of hits as a function of revitalization

Prop. of hits

0.9
0.7
0.5
0.3

■ Without
▥ 6 months
☐ 12 months

Type of revitalization
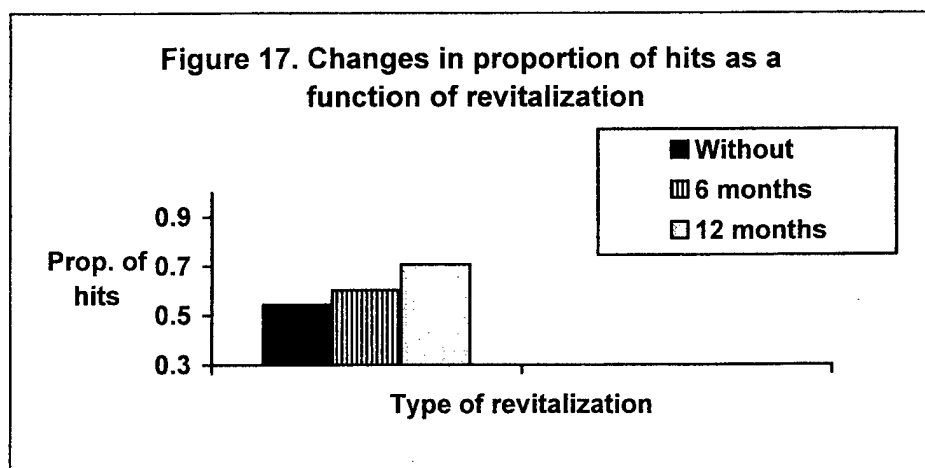
The next figure depicts changes in 'time on target' scores as a function of revitalization. The comparison of the three groups of subjects showed a significant difference [F (2,40) = 4.72, p<0.05]. There was a significant difference between subjects who experienced revitalization and subjects who

had no revitalization [$F(1,40)=9.03$, $p<0.01$]. However, there was no difference between the two groups who had revitalization [$F(1,40)=0.67$, ns].
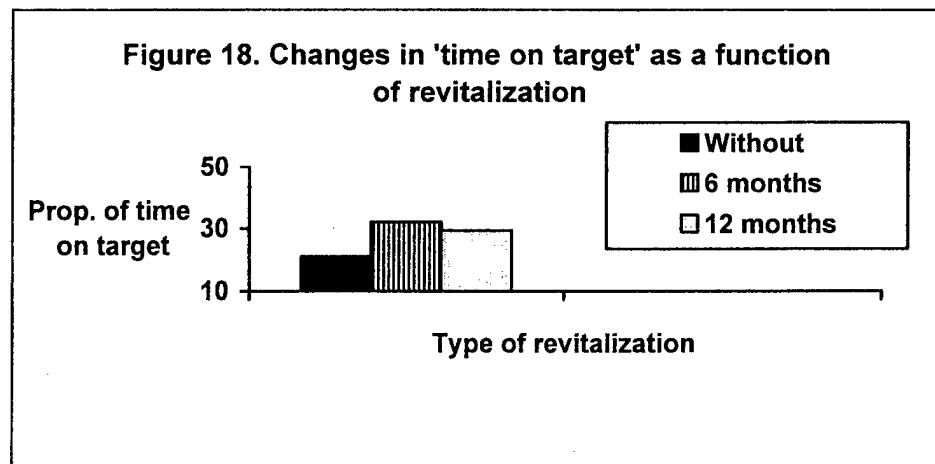
Figure 18. Changes in 'time on target' as a function of revitalization

3. Relationship between knowledge and practice

The order of tests had no effect on performance (all F values were smaller than 1).

Individual differences, original learning and skill decay

We examined predictability of performance after 18 months and skill decay by various individual measures (i.e., KABA, DAPAR, memory performance, and scores in professional courses). This was done separately for each dependent measure (e.g., knowledge scores, various practical measures) by stepwise regressions. Because of the large N (the number of subjects) needed for multiple regression analysis, we merged the data of the three types of skills and analyzed them together. We employed dependent measures which had the same scale for all three skills (i.e., scores in the knowledge tests for all the subjects, and proportion of hits for the missile launchers only).

Table 3 describes the regression summary for predicting scores in the knowledge test after 18 months. The predictor variables can explain 11.52% of the variance. The significant predictors are DAPAR and scores achieved at the professional courses (Beta values are 0.2981 and 0.114, respectively).

Table 3: Prediction of knowledge test scores

	BETA	t value	p level
DAPAR	.298185	4.36233	.000020
Score at the professional course	.114083	1.70101	.090387
Semantic memory	-.089428	-1.34798	.179084
Memory for words	-.087164	-1.31863	.188695

Table 4 describes the regression summary for predicting proportion of hits after 18 months. The predictor variables can explain 12.32 percent of the variance. There was one significant predictor which was 'score at the professional course' (Beta value is 0.237).

Table 4: Prediction of proportion of hits of missile operators

N=122	BETA	t value	p level
Score at the professional course	.237929	2.49784	.013897
KABA	.206706	1.67872	.095899
DAPAR	-.187052	-1.57210	.118651
Memory for order	-.109311	-1.22177	.224272
Memory for pairs	.106371	1.12101	.264599

Measures of motivation

In order to examine whether there was decrement in motivation as a function of the retention interval, we measured (through a questionnaire) the motivation of the participants engaging performance. As mentioned above, the reported results are based on a limited number of participants (i.e., 46). Values of motivation range between 1 and 5 (low and high motivation, respectively). It seems that motivation was high (mean value 3.40, with standard deviation of 0.96). Additionally, we measured the correlation between the motivation scores and the time elapsed since discharge. There was no correlation between motivation measurement and the retention interval (r=0.095, ns).

Self assessment

In order to examine whether participants were able to assess their ability we asked them to evaluate their performance at the end of the day. Table 5 presents correlation coefficients among these evaluations and performance.

Table 5

: Correlations between self assessments and performance

	Tank gunners		Tow operators		Dragon operators	
	Knowledge score	No. of hits	Knowledge score	Pro. Of hits	Knowledge score	Pro. Of hits
Knowledge assessment	0.0379**	0.1350	0.3744**	0.1108	0.2848**	0.1949*
Practical assessment	0.1580*	0.0700	0.0588	0.5804**	0.0639	0.3197**

* significant at p< 0.05, ** significant at p<0.01

Discussion

Following is a brief summary of the results.

There was a decline in performance with time since discharge, both for knowledge and various aspects of shooting performance (e.g., hits, deviations, time on target).

1. Skill refreshing had an effect only on knowledge. This effect appeared regardless of time of revitalization. There was no effect of revitalization on practical aspects of performance.

2. There was no effect of order of tests on performance, neither on knowledge nor on practical aspects of performance.

3. Various background measures predicted performance but explained only a limited percentage of variance. For example, DAPAR predicted knowledge scores significantly but did not predict hit rate. Memory tests did not predict any performance after 18 months.

The most important finding of this study is that both declarative and procedural aspects of performance decline with time. After 18 months performance declines significantly and refreshing soldiers' memory during the period of disuse, by simulator use, cannot prevent such decline. It is not clear if additional practice or more frequent use of simulators might prevent such decline.

Declarative knowledge suffers from disuse but can improve with revitalization. Note that this happens even when revitalization occurs only after 6 months. Moreover, the rate of decay of declarative knowledge is higher than the rate of decay of procedural knowledge. This can be seen when one

compares changes in knowledge tests with changes in hit rates or proportion of hits. For example, Tow operators showed a change from 68.21 to 49.89 in grades of knowledge (25% change) while their hit rate changed from 0.76 to 0.67 (about 10% change). Procedural knowledge does not benefit from revitalization of simulator usage.

It is a common practice in the IDF to have soldiers review their knowledge before using a weapon. We found no support for this practice in this research. If refreshing one's memory for declarative knowledge affects performance, we would expect order of tests to affect the soldiers' performance. However, we did not find any indication of such an effect.

In the Introduction we mentioned differences between discrete and continuous tasks. It is possible to think of tank gun operation as a discrete task while missile launching seems to be more of a continuous task. It was expected that the latter would suffer less from disuse because of the procedural knowledge that is more involved in such a task. However, both types of tasks were affected by disuse. With missile operation even time on target, a measure indicative of tracking ability, declined significantly (both for Dragon and Tow operators). It is possible that the long retention intervals (i.e., 6 months) were responsible for this result.

Various background measures (KABA, DAPAR, memory tests, score at the professional course) which are supposed to be good predictors of performance did not contribute to prediction significantly. One possibility is that the range of these predictors is restricted. However, because operators of these weapons are not selected by these variables, this cannot explain the

present results. With respect to score at the professional course, it is possible that this score was not reliable enough to act as a predictor of performance.

With respect to self assessment, we seem to have replicated early work by Wisher et al. (1991). That is, participants seemed to be able to evaluate their performance correctly. However, Tank gunners assessments were correlated to their knowledge scores only.

References

Anderson, J. R. (1983). The architecture of cognition. Cambridge: Harvard University Press.

Arthur, W. J., Bennet, W. J., Stanush, P. L., & McNelly T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. Human Performance, 11 (1), 57-101.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for spanish Learned in school. Journal of Experimental Psychology: General, 113 ,1-29.

Farr, M. (1987). The long-term retention of knowledge and skills: A cognitive and instructional perspective. New York: Springer-Verlag.

Fisher, D.L., Wisher R. A., & Ranny T. A. (1996). Optimal static and dynamic training Schedules: State model of skill Acquisition. Journal of Mathematical Psychology, 40, 30-47.

Hagman, J. D., & Rose, A. M. (1983). Retention of military tasks: A review. Human Factors, 25(2), 199-213.

Healy, A. F., Fendrich, D. W., Crutcher, R. J., Wittman, W. T., Gesi, A. T., Ericsson, K. A., & Bourne L. E. The long term retention of skills. In. Healy, A. F., Kosslyn, S. M., & Shiffrin, R. M. (Eds.). (1992). From learning processes to cognitive processes: Essays in honor of William K. Este (Vol. 2 , pp. 87-118), Hildsdale, NJ: Erbaum.

Healy, A. F. & Sinclair, G. P. (1996).The long term retention of training and instruction. In Bjork, E. L., & Bjork, R. A. (Eds.), Memory (pp. 525-565). San Diego, California, Academic Press.

Marmie, W. R., & Healy, A. F. (1995). The long-term retention of a complex skill. In Healy, A. F., & Bourne, L. E. (Eds.), Learning and memory of knowledge and skills (30-65), California, Sage publishers.

Mengelkoch, R. F., Adams, J. A., and Gainer, C. A. (1971). The Forgetting of Instrument Flying Skills. Human Factors, 13(5), 397-405.

Spector, A. (1992) Revitalization and retention of military skills within reserve military soldieries. Zichron-Yaa'kov, Israel: The Israeli institution of Military Researches. (in Hebrew)

Wisher, R. A., Sabol, M. A., & Ellis, J. A. (1999). Staying sharp: Retention of

   military knowledge and skills. U.S. Army Research Institute special

   report, 39.

Wisher, R. A., Sabol, M. A., Sukenik, H., and Kern, R. (1991). Individual

   Ready Reserve (IRR) call - up: Skill decay (Research Report 1595).

   Alexandria, VA: U. S. Army Research Institute for Behavioral and

   Social Sciences.